DABOUSSI'S ELEMENTARY PROOF OF THE PRIME NUMBER THEOREM, RECONSIDERED

PAUL POLLACK

Abstract. In 1984, Daboussi gave an "elementary" proof of the prime number theorem whose central actors are smooth numbers (numbers with no large prime factors). We explain a variant of his argument that might be considered even more elementary.

hexagon:2611.00636v1

1. Introduction

Let μ denote the Möbius function, and let $M(x) := \sum_{n \le x} \mu(n)$ be the corresponding summatory function. We will give a simple proof of the following 1897 theorem of von Mangoldt [8].

Theorem 1. *$M(x) = o(x)$, as $x \to \infty$.*

Landau showed that Theorem 1 is "elementarily equivalent" to the prime number theorem (PNT). (His 1899 doctoral thesis proves Theorem 1 from the prime number theorem [6, pp. 6–16]; Landau published the converse implication in 1911 [7, pp. 973–979]. See also the elegant textbook treatment by Shapiro [11, §10.3, pp. 408–414].) Here "equivalent" means that if either statement is assumed, there is a quick proof of the other, significantly shorter than a proof of either statement from scratch. The adverb "elementarily" draws a contrast with the more typical complex function-theoretic methods employed in prime number theory: The quick proofs in question involve only real variables, with Riemann's zeta function never (explicitly) mentioned.

It was quite a sensation when, in 1948, Erdős and Selberg announced that the prime number theorem itself could be established by elementary means [3, 10]. For historical context, see Spencer and Graham's article [12], which reproduces notes by Ernst Straus, and Goldfeld's account [4], which also draws on letters from files Selberg made available to him (cf. the interview [1]). Both Erdős's and Selberg's arguments use Selberg's "symmetry formula"; see [11, eq. (10.4.10), p. 419]. In 1984, Daboussi [2] gave a very different elementary proof of the PNT, using smooth numbers. Further elementary proofs were later given by Hildebrand [5], via a large sieve inequality, and — quite recently — Richter [9], through a general result on the distribution of $\Omega(n)$, the number of prime factors of n counted with multiplicity.

In this note, we present a simplification of Daboussi's argument for Theorem 1. Let $P^+(n)$ denote the largest prime factor of n, adopting the convention that $P^+(1) := 1$. We say n is y-smooth if $P^+(n) \le y$. For $x \ge 1$ and $y \ge 2$, put

$$\Psi(x,y) := \sum_{\substack{n \le x \\ P^+(n) \le y}} 1, \qquad M(x,y) := \sum_{\substack{n \le x \\ P^+(n) \le y}} \mu(n).$$

Set

$$a := \limsup_{x \to \infty} \frac{|M(x)|}{x},$$

2020 *Mathematics Subject Classification.* Primary 11N05; Secondary 11N25.
Key words and phrases. Prime number theorem, Möbius function, smooth numbers, elementary methods.

so that the PNT is equivalent to proving that $a = 0$.

A central component of Daboussi's proof is a comparison between $|M(x,y)|$ and $a\Psi(x,y)$, based on slight variants of the "logarithmic" relations

$$|M(x,y)| \log x \le \sum_{\substack{p \le y \\ p^k \le x}} \log p \, |M(x/p^k, y)| + \int_1^x |M(t,y)| \, \frac{\mathrm{d}t}{t}, \tag{1}$$

$$\Psi(x,y) \log x = \sum_{\substack{p \le y \\ p^k \le x}} \log p \, \Psi(x/p^k, y) + \int_1^x \Psi(t,y) \, \frac{\mathrm{d}t}{t}. \tag{2}$$

(We prove both of these in §4 below.) Daboussi extracts what he needs from (1) and (2) by integrating against the same carefully constructed kernel function; see [2, Lemma 2, p. 162, and §III.3 and §IV, pp. 163–164]. Our proof avoids any need to introduce a special kernel function by working directly with the inequality for $(|M(x,y)| - a\Psi(x,y)) \log x$ that results from subtracting a copies of (2) from (1).

In their lovely book, Tenenbaum and Mendès France describe a variant of Daboussi's argument that, like ours, avoids Daboussi's kernel function [13, Chapter 4, §§5–7, pp. 92–104]. The tradeoff is the need to develop certain ground-floor estimates from the theory of smooth numbers, including that $\Psi(x^u, x)/x^u \approx \rho(u)$ (Dickman's ρ-function) for bounded u (with a key step in their proof being that $|M(x^u, x)|/x^u \lessapprox a\rho(u)$ in the same range). Our arrangement requires no nontrivial estimates on $\Psi(x,y)$.

Of course, we do need *something*. We will make free use of the following two estimates of Mertens, valid for $x \ge 2$:

$$\sum_{p \le x} \frac{\log p}{p} = \log x + O(1), \qquad \prod_{p \le x} \left(1 - \frac{1}{p}\right)^{-1} \ll \log x.$$

For (elementary) proofs, see for example [11, Chapter 9].

Throughout the paper, the variable p is reserved for primes, and the exponent k on a prime power p^k is always assumed to be a positive integer.

Use of LLMs. This is a work of exposition. It is offered as a (hopefully!) human-readable "digestion" of output produced by ChatGPT 6 Astra, as prompted by the author. P. P. takes no credit for the new mathematics but stands with, by, and beside the final manuscript.

Lean formalizations of Theorem 1 and the smooth-number argument presented here, together with verification records and instructions for reproducing the checks, are available at https://github.com/twonth/daboussi-pnt-lean. The formalized results use only Lean's standard axioms of propositional extensionality (`propext`), classical choice (`Classical.choice`), and quotient soundness (`Quot.sound`).

P. P. thanks OpenAI and Anthropic for providing complimentary access through the ChatGPT for Academic Researchers program and the Claude Team plan for scientists, respectively.

Other acknowledgements. P. P. presented this argument in a Fall 2026 graduate course offered at the University of Georgia. He thanks the class, and the university, for allowing him this opportunity to workshop the proof.

2. A Fundamental Lemma and the deduction of Theorem 1

The Möbius function is important in number theory precisely because it provides an elegant, number-theoretically friendly answer to the question of "who is number 1":

$$\mathbf{1}_{n=1} = \sum_{d|n} \mu(d).$$

Letting $x \geq 1$ and summing this identity over $n \leq x$,

$$1 = \sum_{d \leq x} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor = x \sum_{d \leq x} \frac{\mu(d)}{d} - \sum_{d \leq x} \mu(d) \left\{ \frac{x}{d} \right\}.$$

Consequently,

$$\left| \sum_{d \leq x} \frac{\mu(d)}{d} \right| < \frac{1 + \lfloor x \rfloor}{x} \leq 2 \qquad (x \geq 1).$$

Expressing the same sum by partial summation, we have

$$\sum_{n \leq x} \frac{\mu(n)}{n} = \frac{M(x)}{x} + \int_1^x \frac{M(t)}{t} \frac{\mathrm{d}t}{t},$$

and hence

$$\left| \int_1^x \frac{M(t)}{t} \frac{\mathrm{d}t}{t} \right| < 3 \qquad (x \geq 1). \tag{3}$$

Our proof of Theorem 1 combines (3) with the following estimate, whose proof occupies §§3–4.

Lemma 2 (Fundamental Lemma)**.** *As $x \to \infty$,*

$$\int_1^x \left(a - \frac{|M(t)|}{t} \right)^+ \frac{\mathrm{d}t}{t} = o(\log x).$$

(*Here $v^+ = \max\{v, 0\}$ denotes the positive part of v.*)

Roughly speaking, Lemma 2 says that $|M(t)|/t$ spends most of its time close to its limit superior a, with time measured logarithmically (i.e., with respect to the logarithmic measure $\mathrm{d}t/t$).

Proof of Theorem 1, assuming the Fundamental Lemma. Suppose for a contradiction that $a > 0$, and fix $B := \exp(12/a)$. We first use (3) to show that every interval $[X, BX]$, with $X \geq 1$, contains a u for which

$$|M(u)| \leq \frac{a}{2} u. \tag{4}$$

Indeed, if there is no such u, then $M(t)$ is nonzero throughout $[X, BX]$. Since M takes integer values and jumps by at most 1, it has constant sign there. This gives

$$\left| \int_X^{BX} \frac{M(t)}{t} \frac{\mathrm{d}t}{t} \right| = \int_X^{BX} \frac{|M(t)|}{t} \frac{\mathrm{d}t}{t} \geq \frac{a}{2} \log B = 6.$$

But (3), applied at BX and at X, bounds the left-hand side strictly below 6.

Now let $X \geq X_0 := 8/a$, and choose $u \in [X, BX]$ satisfying (4). To obtain a lower bound for the integral in the Fundamental Lemma, we extend this estimate to nearby values of t. For

$u \le t \le u(1+a/8)$, the bound $|\mu(n)| \le 1$ gives

$$\begin{aligned} |M(t)| &\le |M(u)| + \Big| \sum_{u<n\le t} \mu(n) \Big| \\ &\le \frac{a}{2}u + (t-u) + 1 \\ &\le \frac{a}{2}u + \frac{a}{8}u + \frac{a}{8}u \le \frac{3a}{4}t, \end{aligned}$$

where we used $1 \le au/8$. The bounds $u \le BX$ and $a \le 1$ also give $u(1+a/8) \le 2BX$, so

$$\int_X^{2BX} \left(a - \frac{|M(t)|}{t}\right)^+ \frac{\mathrm{d}t}{t} \ge \int_u^{u(1+a/8)} \frac{a}{4}\frac{\mathrm{d}t}{t} = \frac{a}{4}\log(1+a/8) > 0.$$

We apply this bound with $X = (2B)^j X_0$, for $j = 0, \dots, N-1$, noting that the corresponding intervals $(X, 2BX]$ are disjoint and contained in $(1, (2B)^N X_0]$. Adding these estimates gives

$$\int_1^{(2B)^N X_0} \left(a - \frac{|M(t)|}{t}\right)^+ \frac{\mathrm{d}t}{t} \ge \left(\frac{a}{4}\log(1+a/8)\right) N.$$

But the Fundamental Lemma implies that our left-hand side is $o(\log((2B)^N X_0)) = o(N)$ as $N \to \infty$. This contradiction forces $a = 0$. □

3. Splitting off bounded prime factors

Recall from the introduction that $\Psi(x,y) = \sum_{n\le x,\, P^+(n)\le y} 1$ while $M(x,y) = \sum_{n\le x,\, P^+(n)\le y} \mu(n)$. Fix $y \ge 2$, and put

$$P = \prod_{p\le y} p.$$

We have $|M(t,y)| \le \Psi(t,y)$ and, since every positive integer at most y is y-smooth,

$$M(t,y) = M(t), \qquad \Psi(t,y) = \lfloor t \rfloor \qquad (1 \le t \le y).$$

At the other end of the range, only the squarefree y-smooth integers contribute to $M(t,y)$, and these are precisely the divisors of P. Thus, for $t \ge P$,

$$M(t,y) = \sum_{d\mid P} \mu(d) = 0. \tag{5}$$

Following Daboussi, we compare $|M(x,y)|$ with $a\Psi(x,y)$. The following estimate gets us started.

Proposition 3. *For every $y \ge 2$,*

$$\int_1^\infty \left(\frac{|M(t,y)|}{t} - a\frac{\Psi(t,y)}{t}\right)\frac{\mathrm{d}t}{t} \ge 0. \tag{6}$$

Proof. Each positive integer admits a unique decomposition in the form $n = bc$, where $P^+(b) \le y$ and $(c,P) = 1$. Since $(b,c) = 1$, the multiplicativity of μ gives

$$M(x) = \sum_{\substack{c\le x \\ (c,P)=1}} \mu(c) M(x/c, y).$$

The terms with $c \le x/P$ vanish by (5). Taking absolute values in the remaining terms and grouping according to the value of $j := \lfloor x/c \rfloor < P$, we obtain

$$|M(x)| \le \sum_{j=1}^{P-1} |M(j,y)| \sum_{\substack{x/(j+1) < c \le x/j \\ (c,P)=1}} 1.$$

View y as fixed. Since every block of P consecutive integers c contains precisely $\varphi(P)$ that satisfy $(c,P)=1$,

$$\sum_{\substack{x/(j+1)<c\le x/j\\(c,P)=1}} 1 = x\frac{\varphi(P)}{P}\left(\frac{1}{j}-\frac{1}{j+1}\right)+O_y(1).$$

Substituting this into our bound for $|M(x)|$, dividing by x, and letting $x\to\infty$, we find that

$$\begin{aligned} a &\le \frac{\varphi(P)}{P}\sum_{j=1}^{P-1}|M(j,y)|\left(\frac{1}{j}-\frac{1}{j+1}\right)\\ &= \frac{\varphi(P)}{P}\int_1^\infty \frac{|M(t,y)|}{t}\,\frac{\mathrm{d}t}{t}. \end{aligned}$$

(For the final equality, we use that $M(t,y)=M(j,y)$ on $[j,j+1)$, that $M(t,y)=0$ for $t\ge P$, and that $\int_j^{j+1} 1/t\,\mathrm{d}t/t = 1/j-1/(j+1)$.) Rearranging gives

$$\int_1^\infty \frac{|M(t,y)|}{t}\,\frac{\mathrm{d}t}{t} \ge a\frac{P}{\varphi(P)} = a\prod_{p\le y}(1-1/p)^{-1}. \tag{7}$$

The product on the right also occurs in the integral of $\Psi(t,y)/t$. In fact,

$$\int_1^\infty \frac{\Psi(t,y)}{t}\,\frac{\mathrm{d}t}{t} = \sum_{P^+(n)\le y}\frac{1}{n} = \prod_{p\le y}(1-1/p)^{-1}. \tag{8}$$

To see this, observe that by partial summation,

$$\sum_{\substack{n\le X\\P^+(n)\le y}}\frac{1}{n} = \frac{\Psi(X,y)}{X}+\int_1^X \frac{\Psi(t,y)}{t}\,\frac{\mathrm{d}t}{t}. \tag{9}$$

For every positive integer $n\le X$, and every prime p, the exponent of p in n is at most $\log X/\log 2$. Hence (keeping in mind that y is fixed),

$$\Psi(X,y)\le\left(1+\frac{\log X}{\log 2}\right)^{\pi(y)} = o(X).$$

Sending X to infinity in (9) yields the first equality in (8), while the second follows by expanding the Euler product. Finally, the proposition itself follows by comparing (7) and (8). □

To apply Proposition 3 to the Fundamental Lemma, we separate its integrand into positive and negative parts. Define "excess" and "deficit" functions by

$$E_y(t)=\left(\frac{|M(t,y)|}{t}-a\frac{\Psi(t,y)}{t}\right)^+, \qquad D_y(t)=\left(a\frac{\Psi(t,y)}{t}-\frac{|M(t,y)|}{t}\right)^+.$$

The integrand in (6) is $E_y(t)-D_y(t)$, so that Proposition 3 can be restated as the assertion that

$$\int_1^\infty D_y(t)\,\frac{\mathrm{d}t}{t} \le \int_1^\infty E_y(t)\,\frac{\mathrm{d}t}{t}. \tag{10}$$

That is, the total excess of $|M(t,y)|/t$ over $a\Psi(t,y)/t$ is at least the total deficit, both measured with respect to $\mathrm{d}t/t$.

When $1 \le t \le y$, the deficit $D_y(t)$ is closely related to the integrand in the Fundamental Lemma:

$$\begin{aligned} D_y(t) &= \left(a\frac{\lfloor t \rfloor}{t} - \frac{|M(t)|}{t}\right)^+ \\ &= \left(a - \frac{|M(t)|}{t} - a\frac{\{t\}}{t}\right)^+ \ge \left(a - \frac{|M(t)|}{t}\right)^+ - \frac{a}{t}. \end{aligned}$$

(We use here that $(v-h)^+ \ge v^+ - h$ for $h \ge 0$.) Integrating with respect to $\mathrm{d}t/t$, and noting that $\int_1^y a/t\,\mathrm{d}t/t \le a$, we obtain

$$\int_1^y \left(a - \frac{|M(t)|}{t}\right)^+ \frac{\mathrm{d}t}{t} \le \int_1^y D_y(t)\,\frac{\mathrm{d}t}{t} + a \le \int_1^\infty E_y(t)\,\frac{\mathrm{d}t}{t} + a,$$

using (10) in the last step. We have therefore reduced the Fundamental Lemma to proving that, as $y \to \infty$,

$$\int_1^\infty E_y(t)\,\frac{\mathrm{d}t}{t} = o(\log y). \tag{11}$$

4. Making less of the excess

Proof of Lemma 2. To prove (11), we start with the trivial bounds

$$0 \le E_y(t) \le \frac{|M(t,y)|}{t} \le \frac{\Psi(t,y)}{t} \le 1.$$

Together with (8) and Mertens' estimate, this gives

$$\int_1^\infty E_y(t)\,\frac{\mathrm{d}t}{t} \le \prod_{p \le y}(1-1/p)^{-1} \ll \log y. \tag{12}$$

So the remaining fight is over the case of a single letter: We must improve $O(\log y)$ to $o(\log y)$.

We split the integral at y. The required estimate for $\int_1^y E_y(t)\,\mathrm{d}t/t$ follows from the definition of a. Indeed, given $\varepsilon > 0$, choose $T \ge 1$ so that $|M(t)|/t \le a + \varepsilon$ for $t \ge T$. Then, for $T \le t \le y$,

$$E_y(t) = \left(\frac{|M(t)|}{t} - a + a\frac{\{t\}}{t}\right)^+ \le \varepsilon + \frac{a}{t}.$$

Integrating from 1 to y, using $E_y(t) \le 1$ below T, we get that

$$\int_1^y E_y(t)\,\frac{\mathrm{d}t}{t} \le \log T + \varepsilon \log(y/T) + a$$

for all $y \ge T$. Hence,

$$\limsup_{y\to\infty} \frac{1}{\log y}\int_1^y E_y(t)\,\frac{\mathrm{d}t}{t} \le \varepsilon.$$

As $\varepsilon > 0$ is arbitrary,

$$\int_1^y E_y(t)\,\frac{\mathrm{d}t}{t} = o(\log y). \tag{13}$$

To handle the remaining range (y, ∞), we need the "logarithmic" relations quoted in the introduction. We begin their derivation with some simple consequences of Möbius inversion. Writing Λ for the von Mangoldt function and starting from

$$\log n = \sum_{p^k \mid n} \log p = \sum_{d \mid n} \Lambda(d), \tag{14}$$

we find that

$$\begin{aligned}\Lambda(n) &= \sum_{d|n} \mu(d) \log (n/d) \\ &= \log n \sum_{d|n} \mu(d) - \sum_{d|n} \mu(d) \log d \\ &= -\sum_{d|n} \mu(d) \log d.\end{aligned}$$

Inverting this last formula for $\Lambda(n)$ gives

$$-\mu(n) \log n = \sum_{p^k|n} \log p \, \mu(n/p^k). \tag{15}$$

Summing (15) over y-smooth integers $n \le x$, and reversing the order of summation,

$$\begin{aligned}-\sum_{\substack{n\le x \\ P^+(n)\le y}} \mu(n) \log n &= \sum_{\substack{p\le y \\ p^k \le x}} \log p \sum_{\substack{m \le x/p^k \\ P^+(m)\le y}} \mu(m) \\ &= \sum_{\substack{p\le y \\ p^k \le x}} \log p \, M(x/p^k, y).\end{aligned}$$

On the other hand, partial summation gives

$$\sum_{\substack{n\le x \\ P^+(n)\le y}} \mu(n) \log n = M(x,y) \log x - \int_1^x M(t,y) \, \frac{\mathrm{d}t}{t}.$$

Substituting this into the immediately preceding display and rearranging,

$$M(x,y) \log x = -\sum_{\substack{p\le y \\ p^k \le x}} \log p \, M(x/p^k, y) + \int_1^x M(t,y) \, \frac{\mathrm{d}t}{t}. \tag{16}$$

Carrying out the same calculation with (14) replacing (15) gives

$$\Psi(x,y) \log x = \sum_{\substack{p\le y \\ p^k \le x}} \log p \, \Psi(x/p^k, y) + \int_1^x \Psi(t,y) \, \frac{\mathrm{d}t}{t}, \tag{17}$$

which is precisely the identity (2) from the introduction.

Taking absolute values in (16) and appealing to the triangle inequality,

$$|M(x,y)| \log x \le \sum_{\substack{p\le y \\ p^k \le x}} \log p \, |M(x/p^k, y)| + \int_1^x |M(t,y)| \, \frac{\mathrm{d}t}{t}, \tag{18}$$

and so we also have (1).

Subtracting a times (17) from (18), we obtain

$$\begin{aligned}(|M(x,y)| - a\Psi(x,y)) \log x \le \sum_{\substack{p\le y \\ p^k \le x}} &\log p (|M(x/p^k, y)| - a\Psi(x/p^k, y)) \\ &+ \int_1^x (|M(t,y)| - a\Psi(t,y)) \, \frac{\mathrm{d}t}{t}.\end{aligned}$$

Replacing each difference on the right by its positive part only increases the right-hand side and so preserves the inequality. Since the resulting right-hand side is nonnegative, we may also replace the left-hand side by its positive part. Noting that $(|M(t,y)| - a\Psi(t,y))^+ = tE_y(t)$, we deduce after dividing by x that

$$E_y(x)\log x \le \sum_{\substack{p\le y\\ p^k\le x}} \frac{\log p}{p^k} E_y(x/p^k) + \int_1^x \frac{t}{x} E_y(t)\,\frac{\mathrm{d}t}{t}. \tag{19}$$

We integrate (19) with respect to $\mathrm{d}x/x$, over $[1,\infty)$. For each prime power p^k, substituting $x = p^k t$ gives

$$\int_{p^k}^\infty \frac{\log p}{p^k} E_y(x/p^k)\,\frac{\mathrm{d}x}{x} = \frac{\log p}{p^k}\int_1^\infty E_y(t)\,\frac{\mathrm{d}t}{t}.$$

For the remaining term, we reverse the order of integration:

$$\begin{aligned}\int_1^\infty \left(\int_1^x \frac{t}{x} E_y(t)\,\frac{\mathrm{d}t}{t}\right)\frac{\mathrm{d}x}{x} &= \int_1^\infty E_y(t)\left(\int_t^\infty \frac{t}{x}\,\frac{\mathrm{d}x}{x}\right)\frac{\mathrm{d}t}{t}\\ &= \int_1^\infty E_y(t)\,\frac{\mathrm{d}t}{t}.\end{aligned}$$

(Here all manipulations of sums and integrals are justified by nonnegativity.) Combining the two contributions,

$$\int_1^\infty E_y(t)\log t\,\frac{\mathrm{d}t}{t} \le \left(1 + \sum_{p\le y}\sum_{k\ge 1}\frac{\log p}{p^k}\right)\int_1^\infty E_y(t)\,\frac{\mathrm{d}t}{t}.$$

The coefficient on the right is estimated using Mertens' theorem:

$$\sum_{p\le y}\sum_{k\ge 1}\frac{\log p}{p^k} = \sum_{p\le y}\frac{\log p}{p} + \sum_{p\le y}\frac{\log p}{p(p-1)} = \log y + O(1).$$

Hence, for some absolute constant $C > 0$,

$$\int_1^\infty E_y(t)\log t\,\frac{\mathrm{d}t}{t} \le (\log y + C)\int_1^\infty E_y(t)\,\frac{\mathrm{d}t}{t}.$$

Rearranging and applying (12), we find that

$$\int_1^\infty E_y(t)\log(t/y)\,\frac{\mathrm{d}t}{t} \le C\int_1^\infty E_y(t)\,\frac{\mathrm{d}t}{t} = O(\log y). \tag{20}$$

With the factor $\log(t/y)$ omitted, we already had an $O(\log y)$ bound; the advantage here is that the new factor is sensitive to *where* the excess occurs.

Let's mop up. Recall that $\int_1^y E_y(t)\,\mathrm{d}t/t = o(\log y)$. Slightly larger values of t, say $y \le t \le y^{1+\varepsilon}$ with $\varepsilon > 0$ fixed, pose no trouble. Since $E_y(t) \le 1$,

$$\int_y^{y^{1+\varepsilon}} E_y(t)\,\frac{\mathrm{d}t}{t} \le \int_y^{y^{1+\varepsilon}}\frac{\mathrm{d}t}{t} = \varepsilon\log y,$$

which will be acceptable to us upon taking ε small.

It remains to consider t that are substantially larger than y, meaning $t \geq y^{1+\varepsilon}$. Moving the contribution from $[1, y]$ in (20) to the right and using (13), we obtain

$$\begin{aligned}\int_y^\infty E_y(t)\log(t/y)\,\frac{\mathrm{d}t}{t} &\leq \int_1^y E_y(t)\log(y/t)\,\frac{\mathrm{d}t}{t} + O(\log y)\\ &\leq (\log y)\int_1^y E_y(t)\,\frac{\mathrm{d}t}{t} + O(\log y)\\ &= o((\log y)^2). \qquad (21)\end{aligned}$$

On the other hand, $\log(t/y) \geq \varepsilon \log y$ when $t \geq y^{1+\varepsilon}$, so

$$\int_y^\infty E_y(t)\log(t/y)\,\frac{\mathrm{d}t}{t} \geq \int_{y^{1+\varepsilon}}^\infty E_y(t)\log(t/y)\,\frac{\mathrm{d}t}{t} \geq \varepsilon \log y \int_{y^{1+\varepsilon}}^\infty E_y(t)\,\frac{\mathrm{d}t}{t}.$$

Comparing with (21), we find that

$$\int_{y^{1+\varepsilon}}^\infty E_y(t)\,\frac{\mathrm{d}t}{t} = o(\log y).$$

Combining the estimates for the three ranges gives $\int_1^\infty E_y(t)\,\mathrm{d}t/t \leq \varepsilon \log y + o(\log y)$. Dividing by $\log y$ and letting $y \to \infty$, we obtain

$$\limsup_{y\to\infty} \frac{1}{\log y}\int_1^\infty E_y(t)\,\frac{\mathrm{d}t}{t} \leq \varepsilon.$$

But $\varepsilon > 0$ is arbitrary, and so $\int_1^\infty E_y(t)\,\mathrm{d}t/t = o(\log y)$, as desired. This proves (11) and hence the Fundamental Lemma. As shown in §2, it follows that $M(x) = o(x)$. □

References

[1] N. A. Baas and C. F. Skau, *The lord of the numbers, Atle Selberg. On his life and mathematics*, Bull. Amer. Math. Soc. (N.S.) **45** (2008), no. 4, 617–649. An extended version of the interview is available at https://www.math.ntnu.no/Selberg-interview/.

[2] H. Daboussi, *Sur le théorème des nombres premiers*, C. R. Acad. Sci. Paris Sér. I Math. **298** (1984), no. 8, 161–164.

[3] P. Erdős, *On a new method in elementary number theory which leads to an elementary proof of the prime number theorem*, Proc. Nat. Acad. Sci. U.S.A. **35** (1949), 374–384.

[4] D. Goldfeld, *The elementary proof of the prime number theorem: an historical perspective*, in *Number Theory* (New York, 2003), D. Chudnovsky, G. Chudnovsky, and M. Nathanson (eds.), Springer, New York, 2004, 179–192.

[5] A. Hildebrand, *The prime number theorem via the large sieve*, Mathematika **33** (1986), 23–30.

[6] E. Landau, *Neuer Beweis der Gleichung* $\sum_{k=1}^\infty \mu(k)/k = 0$, Inauguraldissertation, Berlin, 1899.

[7] ———, *Über die Äquivalenz zweier Hauptsätze der analytischen Zahlentheorie*, Sitzungsber. Kaiserl. Akad. Wiss. Wien, Math.-Naturwiss. Kl., Abt. IIa **120** (1911), 973–988.

[8] H. von Mangoldt, *Beweis der Gleichung* $\sum_{k=1}^\infty \mu(k)/k = 0$, Sitzungsber. Königl. Preuß. Akad. Wiss. Berlin (1897), 835–852.

[9] F. K. Richter, *A new elementary proof of the Prime Number Theorem*, Bull. London Math. Soc. **53** (2021), 1365–1375.

[10] A. Selberg, *An elementary proof of the prime-number theorem*, Ann. of Math. (2) **50** (1949), 305–313.

[11] H. N. Shapiro, *Introduction to the Theory of Numbers*, John Wiley & Sons, New York, 1983; reprinted by Dover Publications, Mineola, NY, 2008.

[12] J. Spencer and R. Graham, *The elementary proof of the prime number theorem*, Math. Intelligencer **31** (2009), no. 3, 18–23.

[13] G. Tenenbaum and M. Mendès France, *The Prime Numbers and Their Distribution*, translated by P. G. Spain, Student Mathematical Library, vol. 6, American Mathematical Society, Providence, RI, 2000.

Department of Mathematics, University of Georgia, Athens GA 30602

Email address: pollack@uga.edu